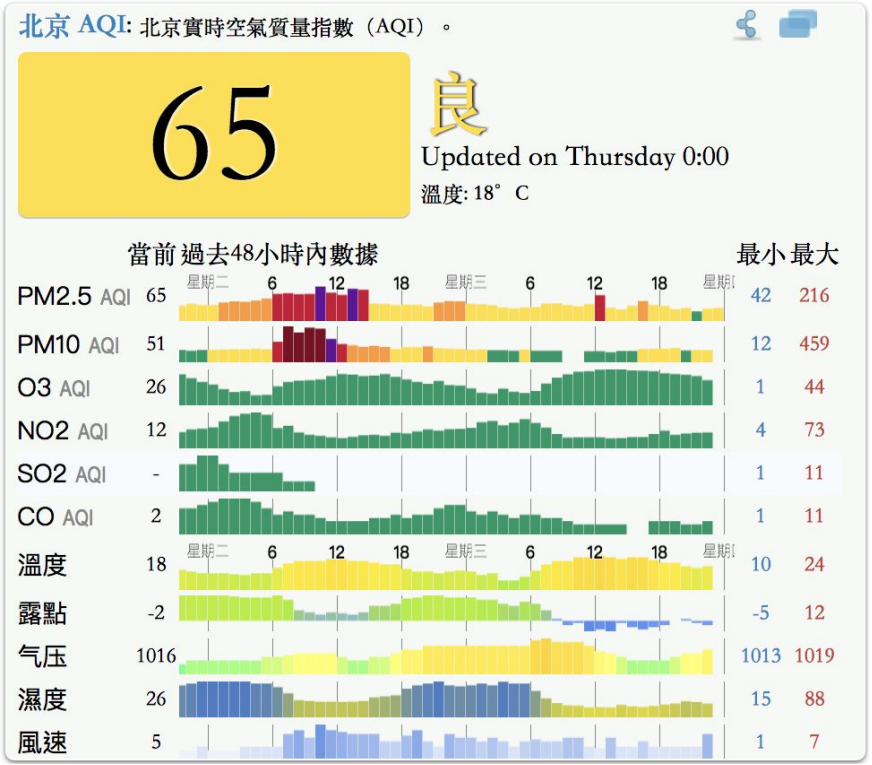
Neural Net
Midterm Report

Reporter: Wei-Yuan Hsu
Advisor: Prof. Hahn-Ming Lee

Problem Definition



Data Exploration

Air Quality Station	Number	Name
Urban	12	Ex. dongsi (東城東四)
Suburban	11	Ex. fangshan_aq 房山良鄉
Reference	7	Ex. dingling_aq 昌評定陵
Traffic	5	Ex. qianmen_aq 前門東大街

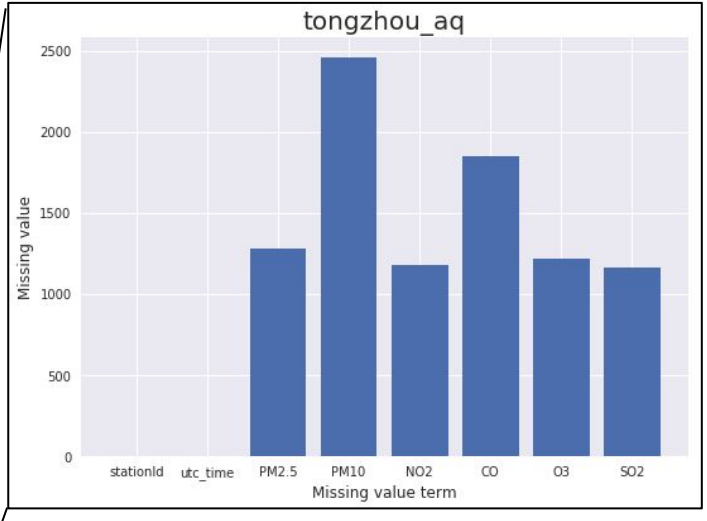
Data Exploration

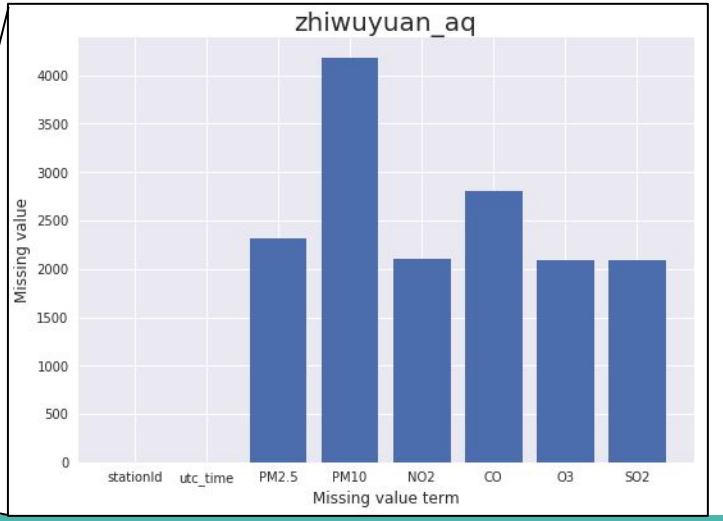
```
In [11]:  1  air_station.size()

Out[11]: stationId
         aotizhongxin_aq      8886
         badaling_aq          8886
         beibuxinqu_aq        8886
         daxing_aq            8886
         dingling_aq          8886
         donggaocun_aq        8886
         dongsi_aq            8886
         dongsihuan_aq        8886
         fangshan_aq          8886
         fengtaihuayuan_aq    8886
         guanyuan_aq          8886
         gucheng_aq           8886
         huairou_aq           8886
         liulihe_aq           8886
         mentougou_aq         8886
         miyun_aq             8886
         miyunshuiku_aq       8886
         nansanhuan_aq        8886
         nongzhanguan_aq      8886
         pingchang_aq         8886
         pinggu_aq            8886
         qianmen_aq           8886
         shunyi_aq            8886
```

Any missing Value?


aotizhongxin_aq
badaling_aq
beibuxinqu_aq
daxing_aq
dingling_aq
donggaocun_aq
dongsi_aq
dongsihuan_aq
fangshan_aq
fengtaihuayuan_aq
guanyuan_aq
gucheng_aq
huairou_aq
liulihe_aq
mentougou_aq
miyun_aq
nansanhuan_aq
pingchang_aq
pinggu_aq
qianmen_aq
shunyi_aq
tiantan_aq
tongzhou_aq
wanliu_aq
wanshouxigong_aq
xizhimenbei_aq
yanqin_aq
yizhuang_aq
yongdingmennei_aq
yongledian_aq
yufa_aq
yungang_aq
zhiwuyuan_aq


tongzhou_aq
Missing value
0
500
1000
1500
2000
2500
stationId
utc_time
PM2.5
PM10
NO2
CO
O3
SO2
Missing value term


zhiwuyuan_aq
Missing value
0
500
1000
1500
2000
2500
3000
3500
4000
stationId
utc_time
PM2.5
PM10
NO2
CO
O3
SO2
Missing value term

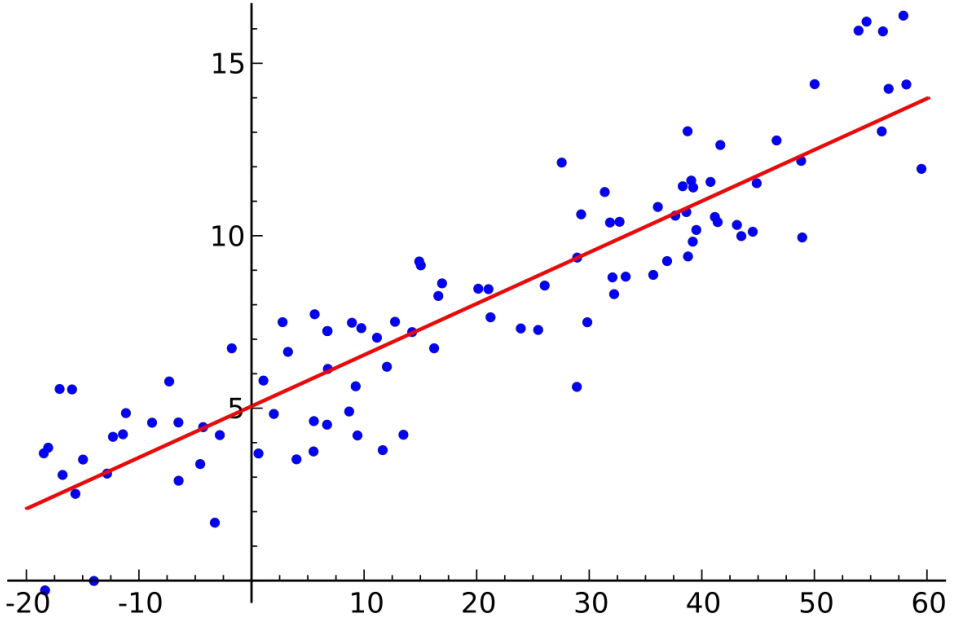
Data Exploration - Missing Value

Method#1	Fill NaN with Value ex. zero
Method#2	Drop Data
Method#3	…

Methodology

Q: What kind of question it is?

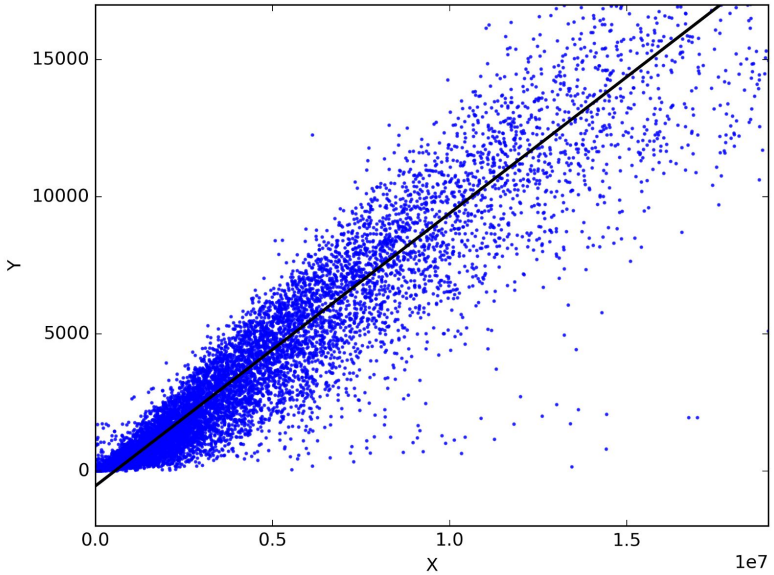
A: It is a typical regression analysis

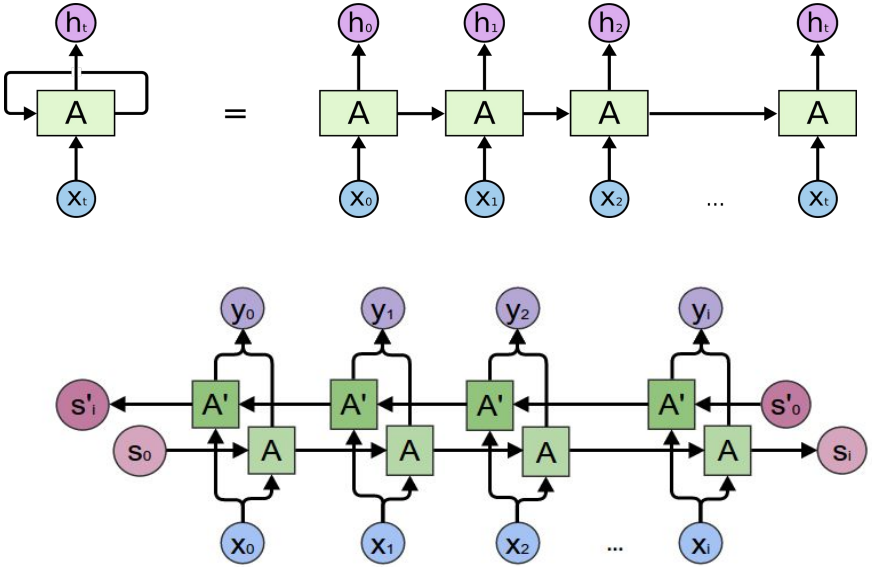


[1] Wiki - Regression analysis, https://en.wikipedia.org/wiki/Regression_analysis

Model

Element	LSTM
Hidden Layer	1
Output Layer	1





[1] ML 101: Linear Regression Tutorial,
https://medium.com/@amarbudhiraja/ml-101-linear-regression-tutorial-1e40e29f1934

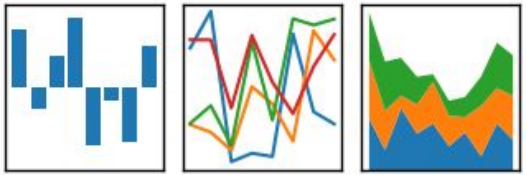
Tools

python™

machine learning in Python

pandas

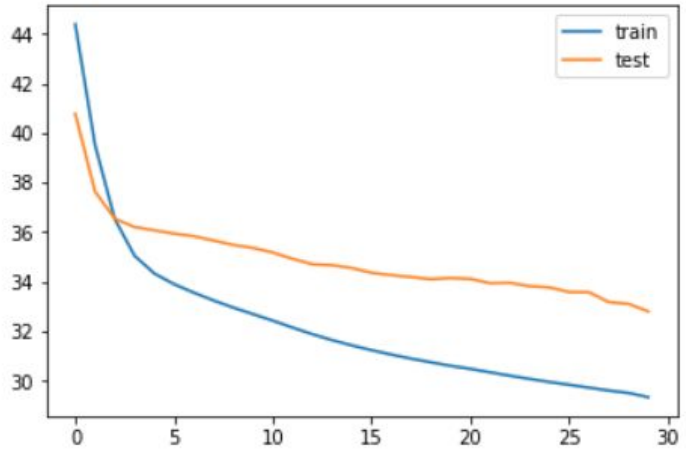
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Keras

Result

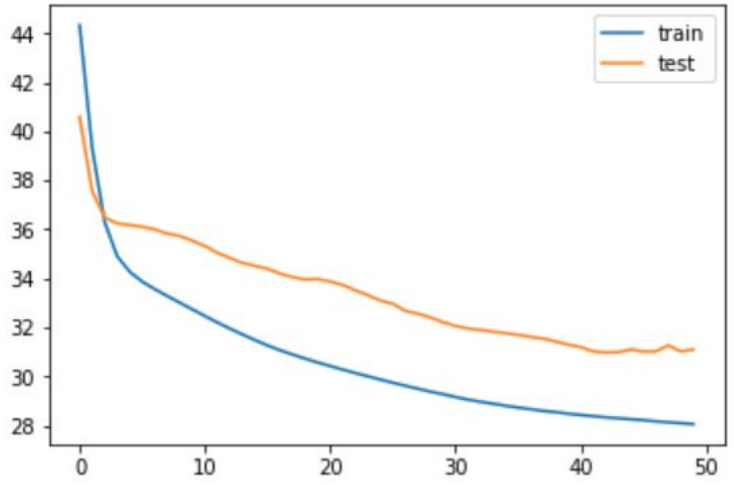
Loss



Epoch

Missing Value	Fill NaN
Epoch	30
Model	RNN

Loss



Epoch

Missing Value	Fill NaN
Epoch	50
Model	RNN

TODO

Topic	Solution
Feature Enginnering	Feature Importance Analysis
Validation Automation	Pipeline
Model	Ensemble / XGBoost